

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1-38. (Canceled)

39. (New) A method of linking data records from a plurality of sources, the method comprising the steps of:

- a) receiving data records from a plurality of data sources;
- b) cleaning the received data records;
- c) matching received data records together according to matching rules for examining the similarity between respective data records, and building a virtual data model, based on matching results, that represents the data records as a map of linked data records, wherein the data records are not merged by the matching step; and
- d) generating a report based on at least the matching results represented by the virtual data model.

40. (New) A method of linking data records according to claim 39 further comprising the step of modifying the matching rules so as to build at least one modified virtual data model.

41. (New) A method of linking data records according to claim 40 wherein the step of modifying the matching rules so as to build a modified virtual data model is performed iteratively.

42. (New) A method of linking data records according to claim 39 further comprising the step of splitting the received data records into a plurality of respective data types.

43. (New) A method of linking data records according to claim 42 wherein the splitting step is performed before the cleaning step.

44. (New) A method of linking data records according to claim 39 wherein an audit trail is maintained for any changes made to the received data records during the cleaning step.

45. (New) A method of linking data records according to claim 44 wherein the report is based on at least the matching results represented by the virtual data model and the audit trail.

46. (New) A method of linking data records according to claim 39 wherein the report is indicative of the quality of the data records from the plurality of sources.

47. (New) A method of linking data records according to claim 39 comprising the further step of standardizing a format of the received data records before the cleaning step.

48. (New) A method of linking data records according to claim 42 in which data records have multiple data types and in which the cleaning step is carried out independently in respect of some or all of the data types.

49. (New) A method of linking data records according to claim 48 in which the data types comprise names and addresses, and the cleaning step is applied to names and addresses included in the received data records.

50. (New) A method of linking data records according to claim 48 in which the data types include at least one of: dates; reference numbers; telephone numbers; e-mail addresses and

cleaning is carried out in respect of any one or any combination of these other data types.

51. (New) A method of linking data records according to claim 39 in which the cleaning step comprises a step of standardizing the data records against a predetermined standard.

52. (New) A method of linking data records according to claim 51 in which the predetermined standard comprises a predetermined list.

53. (New) A method of linking data records according to claim 52 which is such as to allow a user to select at least one list against which data records are to be standardized.

54. (New) A method of linking data records according to claim 39 in which the cleaning step comprises standardizing the data records through an application of rules.

55. (New) A method of linking data records according to claim 54 which is such as to allow a user to select at least one rule which is applied to the data records in the cleaning step.

56. (New) A method of linking data records according to claim 54 in which the rules are used to at least one of: change the data records to a standardized form, correct data records, and complete data records.

57. (New) A method of linking data records according to claim 39 in which standardization against a list is performed in combination with standardization through rules.

58. (New) A method of linking data records according to claim 39 in which the cleaning step comprises an automated cleaning process which is intelligent such that it learns from decisions made by human intervention.

59. (New) A method of linking data records according to claim 39 wherein the matching step matches data records which relate to a common entity and which originate from respective distinct data sources.

60. (New) A method of linking data records according to claim 59 in which the step of matching data records comprises the step of comparing a plurality of data items in respective data records to decide whether the data records relate to a common entity.

61. (New) A method of linking data records according to claim 60 in which at least one threshold level of similarity between data items is specified, such that the threshold must be met or exceeded before a match is determined.

62. (New) A method of linking data records according to claim 59 in which matching rules specify a plurality of matching criteria at least one of which must be met before a match can be determined.

63. (New) A method of linking data records according to claim 62 in which each matching criterion identifies at least one predetermined type of data item and at least one similarity threshold.

64. (New) A method of linking data records according to claim 39 wherein an audit trail is maintained so as to keep a record of matches made in the matching step.

65. (New) A method of linking data records according to claim 39 in which an output of the matching process is used to modify the cleaning step.

66. (New) A method of linking data records according to claim 39 in which the method comprises the further step of de-duplication of data records.

67. (New) A method of linking data records according to claim 66 in which an audit trail is maintained so as to keep a record of changes made to the data records in the de-duplication step.

68. (New) A method of linking data records according to claim 39 in which the cleaning step is performed iteratively.

69. (New) A method of linking data records according to claim 66 in which the de-duplication step is performed iteratively.

70. (New) A method of linking data records according to claim 42 wherein an audit trail is maintained comprising any changes made to the data records during the cleaning step, and wherein the audit trail is maintained for each data type so that there are audit entries in respect of every part of every data record that has been modified.

71. (New) A method of linking data records according to claim 44 in which the audit trail comprises a measure of the quality of the data records.

72. (New) A computer program product comprising at least one data carrier carrying a computer program comprising code portions that when loaded and run on a computer cause the computer to carry out a method of:

- a) receiving data records from a plurality of data sources;
- b) cleaning the received data records;
- c) matching received data records together according to matching rules for examining the similarity between respective data records, and building a virtual data model, based on matching results, that represents the data records as a map of linked data records, wherein the data records are not merged by the matching step; and
- d) generating a report based on at least the matching results represented by the virtual data model.

73. (New) A method of generating a virtual data model comprising the steps of:

- a) receiving data from a plurality of data sources;
- b) cleaning the received data records; and
- c) linking received data records together according to matching rules for examining the similarity between respective data records, and building a virtual data model, based on matching results, that represents the data records as a map of linked data records.

74. (New) A method of linking data records from a plurality of sources comprising the steps of:

- a) receiving data records from a plurality of data sources;
- b) cleaning the received data records, wherein an audit trail is maintained for any changes made to the received data records during the cleaning step;
- c) matching received data records together according to matching rules for examining the similarity between respective data records, and building a virtual data model while maintaining data records as unmerged entities; and
- e) generating a report based on at least the virtual data model and the audit trail, the report indicating the quality of the data records from a plurality of sources.

75. (New) A method of aggregating data comprising the steps of:

- a) receiving data from a plurality of sources;
- b) cleaning the received data, whilst maintaining an audit trail of any changes made to the data in the cleaning step;
- c) creating a data set comprising the cleaned data and the audit trail; and
- d) generating output data using said data set.

76. (New) Apparatus arranged under the control of software for linking data records from a plurality of sources by:

- a) receiving data records from a plurality of data sources;

- b) cleaning the received data records;
- c) matching received data records together according to matching rules for examining the similarity between respective data records, and building a virtual data model, based on matching results, that represents the data records as a map of linked data records, wherein the data records are not merged by the matching step; and
- d) generating a report based on at least the matching results represented by the virtual data model.

77. (New) Apparatus according to claim 76 which is arranged to output a query notification when unable to automatically clean a data item.

78. (New) Apparatus according to claim 77 which is arranged to, allow input of a decision to resolve the query, and complete the cleaning step for that data item based on that decision.

79. (New) Apparatus according to claim 76 which is arranged to learn from a decision input to resolve a query to aid in the cleaning of future data items.

80. (New) Apparatus which is arranged to produce a computer program product comprising at least one data carrier carrying a computer program comprising code portions that when loaded and run on a computer, arrange the computer as an apparatus to:

- a) receive data records from a plurality of data sources;
- b) clean the received data records;
- c) match received data records together according to matching rules for examining the similarity between respective data records, and building a virtual data model, based on matching results, that represents the data records as a map of linked data records, wherein the data records are not merged by the matching step; and
- d) generate a report based on at least the matching results represented by the virtual data model.

81. (New) A method of generating a virtual data model representing data held by an organization in a plurality of distinct data sources comprising the steps of:

- a) receiving data from the plurality of data sources;
- b) cleaning the received data, whilst maintaining an audit trail of any changes made to the data in the cleaning step;
- c) creating a virtual data model, comprising the cleaned data, linkages between the cleaned data, and the audit trail.